Metaphysical Neutrality: Untangling Entanglement

Physicalism, defined here as the substance monist view that nothing exists other than matter, is often given a structural advantage that it hasn't earned evidentially. A common claim is that science supports a mechanistic understanding of reality. This claim isn't nearly as convincing as it first appears. That is not to say that science supports another view of reality. It simply means we ought to drop the assumption that science plays for team materialism and realise that the ultimate questions still compete on a level playing field.

Physics operates under a principle of methodological naturalism — the principle that science ought to prioritise physical causes for physical effects over any extra-physical explanations. This is the appropriate methodology for studying physical phenomena, but it becomes a metaphysical assumption when extended to claim that only physical causes exist. That additional step goes beyond the evidence. By definition, metaphysics is beyond the physical sciences.

The method itself is predisposed toward physical explanations. You could comb a beach with a metal detector without finding any buried diamonds, even if the beach is littered with them. You might, by pure chance, come across one or two. I think this is what quantum mechanics reveals — glimpses of something the method wasn't designed to detect.

There is no telescope that can observe what lies beyond the edges of the universe, and no microscope that can objectively observe inner experience. The tools of science are not designed to look for such things. That is not a criticism, just a description.

Yet quantum mechanics, despite operating entirely within the physical sciences, has generated findings that sit uneasily with purely physicalist interpretations. Methodological naturalism is a commitment adopted prior to the evidence, not a conclusion reached through it. It has shaped interpretations of quantum mechanics in ways that favour physicalism without decisive evidential support. Quantum phenomena can be accommodated at least as readily — and in some cases more naturally — by perspectives that treat consciousness as fundamental rather than emergent.

Quantum entanglement — what Einstein called "spooky action at a distance" — offers a first glimpse of this. When two particles interact they can become entangled, remaining directly

correlated regardless of the distance between them. Measure one and the other responds instantaneously, even across vast distances. They are not transmitting information across space in any conventional sense. They remain in fundamental relation to one another in a way that classical physics cannot account for. The assumption of a world of distinct, independently existing objects interacting through local causes struggles to accommodate this naturally. It is, however, exactly what we would expect if connection is more fundamental than separation.

The same is true of the observer effect. In the 1920's and 1930's the pioneers of quantum science observed particles in an indeterminate state of 'superposition' prior to observation. The particles did not appear to occupy a specific location if no observation was taking place. They were genuinely indeterminate — not in an unknown location, but in no definite location at all — until observation resolved them into a specific state.

The particles of classical physics turned out not to be as fundamental as had been assumed. The language shifted to describe what was actually observed: a superpositioned wavefunction rather than a particle with definite properties. Observation seemed to collapse this wavefunction, forcing the particle to assume a specific location. One interpretation held that consciousness had triggered the collapse. This interpretation requires that consciousness be fundamental rather than emergent — a possibility methodological naturalism is structurally predisposed to rule out.

A key question remains unresolved: was the wavefunction reacting to consciousness, or was it the act of measurement that determined the outcome? Did it assume a position when the data was collected by an instrument, or would it still be 'waiting' for a conscious observer to look at it? There is no way of checking without introducing the variable of consciousness itself, which would defeat the purpose. The question cannot be settled empirically.

The Von Neumann Chain offered an intriguing mathematical explanation for the observer effect. Quantum mechanics applies to everything made of atoms — which is everything. If the original particle is in superposition, then the detector that interacts with it becomes entangled with that superposition and inherits its indefiniteness. The same applies to any recording device, to neural signals, and ultimately to the brain state itself. No physical process within the chain can force a resolution — the Schrödinger equation simply carries the superposition forward through each link. John von Neumann himself concluded that

consciousness was the only element capable of terminating this regress. Not because consciousness is an obvious or mystical answer, but because it is the only element in the sequence that does not behave according to the physical dynamics that keep everything else in superposition. If the chain requires something outside the purely physical to bring about a definite outcome, then consciousness becomes not a speculative add-on but the only candidate that fits the role.

Consciousness appears as a plausible candidate for what breaks the chain, not because it is the obvious answer, but because it is the only element in the sequence that plausibly operates beyond the physical processes the equation describes. If the chain needs something beyond the purely physical to resolve the equation, then the possibility that consciousness operates beyond the purely physical becomes more than speculation — it becomes the best available explanation. It explains the data better than anything else.

According to physicalist metaphysics, consciousness is merely an emergent feature of physical processes, entirely reducible to brain activity. If so, consciousness is just another physical link in the entanglement chain with no special status. But if that's the case, nothing remains that can plausibly resolve the superposition.

Methodological naturalism is therefore predisposed toward finding an alternative explanation — one that stays within the physical. This is a consequence of the method, not the evidence. The metal detector looks for metal, even when the beach is littered with diamonds.

Since we observe particles occupying definite locations — as classical physics requires — an alternative explanation was needed. Since the 1980s the prevailing interpretation among physicists has been the principle of decoherence.

The particle in superposition becomes entangled with its environment — air molecules, photons, electromagnetic fields. The superposition spreads across this massive entangled system. To detect that the particle is still in superposition, we would need to observe quantum interference between all those environmental particles simultaneously, which is practically impossible. Superposition remains, just "diluted" across so many particles that we can no longer observe it. The particle appears to be in one place because we have lost access to the quantum information showing it is in multiple places.

For practical purposes, decoherence explains everything physicists need to predict and manipulate quantum systems. It accounts for why macroscopic objects behave classically and why we never observe superposition at everyday scales. This is no small achievement — it solves the practical problem comprehensively. But it explains why we do not observe superposition continuing, not why we observe a definite position at all. If the particle never actually collapsed — if it merely became unobservably spread out — then what determines which position we see when we look? Decoherence cannot fill this ontological gap. It only describes the conditions under which collapse becomes undetectable, not the conditions under which collapse actually occurs. It asserts that the wavefunction technically continues in superposition indefinitely, but behaves classically for all practical purposes. This is an epistemological solution (we can't observe it anymore) rather than an ontological one (it actually collapsed into one state).

Has the wavefunction collapsed at the moment of instrument detection, will it collapse at the moment of human observation, or does it never collapse? Decoherence alone cannot answer this. And yet we observe particles in specific locations. A cat in our experienced reality is either dead or alive, not both. If nothing ever actually collapses into definite states, then causation, persistence, and macroscopic reality itself become metaphysically mysterious. Decoherence makes the observations fit classical expectations by pushing the ontological question outside the remit of empirical investigation. Physicists need not resolve this because it is a philosophical question, not a scientific one. It can only be answered interpretively.

And so physicists venture into philosophical territory to defend the physicalist interpretation. At that point they are no longer speaking as empirical scientists but as philosophers — and their interpretations carry no more evidential weight than any other philosophical position. The commitment to physicalism is not irrational. It has led to enormous technological success, unified explanations across scientific domains, and avoided the conceptual tangles of substance dualism. For four centuries it has been extraordinarily productive. But productivity in one domain doesn't settle metaphysical questions in another. The question here is whether quantum mechanics specifically vindicates physicalism, and on that narrower question the evidence is far less decisive.

Physicists may invoke parsimony: the physical explanation is simpler because it doesn't add an extra unexplained ingredient. That would be convincing if the two theories held the same explanatory status. They do not. Decoherence doesn't solve the measurement problem — it

relocates it. Instead of asking 'why does collapse occur?' it says 'superpositions become unobservable.' This trades an ontological question for an epistemological convenience.

Consciousness-collapse, by contrast, offers a genuine resolution. If consciousness is what resolves superposition into actuality, and if every physical element in the chain remains governed by quantum mechanics and thus stays in superposition, then consciousness must be operating beyond the purely physical — otherwise it too would remain in superposition and resolve nothing. Consciousness-collapse has its own explanatory gap — we have no mechanism for exactly how consciousness resolves superposition — but this gap is no larger than those in the physical alternatives, and at least points to something we know exists.

A common objection to consciousness-collapse is that a non-physical mind could not influence the physical world without violating conservation of energy. But this objection misunderstands what collapse is. Collapse — in every interpretation — is not an ordinary physical interaction governed by the usual dynamical laws. It lies outside the Schrödinger equation entirely, and therefore outside the domain where conservation theorems straightforwardly apply.

Physicalist collapse theories face exactly the same difficulty. GRW models explicitly introduce energy violations during spontaneous collapse. Pilot-wave theory requires non-local influences that sit uncomfortably with relativistic physics. Many-worlds must preserve conservation across an ever-branching multiverse of unobservable universes. The conservation problem is not unique to consciousness-collapse — it is a problem for collapse itself, regardless of what triggers it.

If consciousness is non-physical, it need not transfer energy or momentum to the system. Selecting which outcome becomes actual from among the physically possible outcomes is not a dynamical push on particles but a resolution of indeterminacy. Conservation laws constrain physical interactions between physical entities. They do not obviously constrain what resolves an indeterminate quantum state into a definite classical one. The objection therefore fails to distinguish consciousness-collapse from its physicalist rivals, all of which face equal or greater tension with conservation principles.

The physicalist attempts to reach the same level of explanatory power in order to preserve physical ontology, but they do so by trading one mystery for another. Everett's 'many worlds' formulation supposes that no wavefunction collapse is necessary and that all possibilities

continue, branching out into potentially limitless worlds at all moments and in all directions. Many Worlds has genuine appeal — it requires no modification to quantum mechanics and takes the Schrödinger equation completely seriously, allowing it to govern everything without exception. But this comes at a considerable metaphysical cost. It postulates infinite unobservable universes to avoid one non-physical entity, arguably violating Occam's razor more spectacularly than consciousness-collapse does. It also offers no explanation for why we experience only one branch, or how that branch is selected from the infinite alternatives.

Physicists have proposed purely physical alternatives to consciousness-collapse. The most developed is spontaneous collapse theory, proposed by Ghirardi, Rimini, and Weber in 1986. According to this view, the wavefunction collapses spontaneously at random intervals without any observer or measurement. For individual particles this occurs extremely rarely, but macroscopic objects contain billions of particles, so at least one undergoes collapse almost instantly, dragging the entire entangled system with it. This explains why we never observe macroscopic superpositions without invoking consciousness. However, it achieves this by modifying the Schrödinger equation with an arbitrary collapse threshold. There is no explanation for why collapse happens, what mechanism triggers it, or why it occurs at that particular rate. It simply asserts that collapse is a brute physical fact — relocating the mystery rather than resolving it.

The pilot wave interpretation, originally proposed by de Broglie in 1927 and developed by David Bohm in 1952, offers another alternative. According to this view, particles always have definite positions, guided by a "pilot wave" that determines their motion. There is no collapse — the wavefunction is real but supplementary, steering particles along deterministic paths. This preserves classical definiteness without invoking consciousness. However, it requires non-local hidden variables that can never be measured. The pilot wave instantly affects particle motion across arbitrary distances, requiring non-local action. It posits a two-layer reality — observable particles plus unobservable guiding fields — that goes well beyond what the evidence requires. Like spontaneous collapse, it solves the measurement problem by adding unexplained physical machinery rather than working with what we already know exists.

Each of these interpretations makes metaphysical commitments comparable to or greater than consciousness-collapse. The difference is they appear more "scientific" because they stay within physical ontology. In doing so they do not solve the problem more convincingly —

they just relocate the mystery to different unmeasurables. Consciousness-collapse at least has the virtue of pointing to something we know exists (consciousness) rather than postulating things we can never observe (other universes, hidden variables, spontaneous collapse mechanisms).

The physicalist may object: if consciousness collapses the wavefunction, and consciousness is inherently subjective, why do all observers see the same result? The intersubjective coherence of quantum measurements — the fact that different observers always agree on outcomes — is sometimes presented as evidence that something objective and physical must be responsible for collapse, not something as variable as individual consciousness.

But this objection assumes that different consciousnesses would naturally impose different resolutions. There is no reason to expect this. By way of analogy, two thermometers measure the same temperature because they're the same kind of instrument measuring the same phenomenon. Two consciousnesses report the same quantum result for exactly the same reason. Intersubjective coherence does not require a special explanation. It is what happens when two instruments of the same kind interact with the same phenomenon.

A physicalist who proposes that two instances of human consciousness could produce different effects on physical reality is ignoring biology. Consciousness evolved over millions of years in a shared physical environment. Its consistency isn't a coincidence — it's a biological requirement. If two observers collapsed the same wavefunction into different outcomes, could they still coordinate, communicate, or survive? The intersubjective coherence of observation isn't evidence that the observer is physical. It's evidence that the observer is uniform. These are not the same.

Finally a physicalist may point to the success rate of physicalism. If it is just one interpretation among many, why has it been so successful at prediction and technological application? The answer is that predictive power belongs to the mathematics, not the metaphysics. Quantum mechanics is famously described as "shut up and calculate" — the equations work perfectly regardless of interpretation. A physicalist and an idealist both use the same Schrödinger equation to build the same functioning MRI machine. The prediction comes from the math, not from the claim that matter is all that exists.

We think physicalism is uniquely successful because we've defined success as manipulating the physical world. By that metric, yes, physicalism excels. But if we define success as

explaining why we have conscious experience at all, physicalism has been notably unsuccessful. The predictive power of science is metaphysically neutral. It can be supplemented with a physicalist, idealist, or panpsychist interpretation, but the interpretation isn't what makes the technology work.

None of this is proof of a non-physicalist reality. Such proof, in either direction, may be impossible for the physical sciences. But the evidence significantly weakens the assumption that physics and physicalism go hand in hand, and suggests consciousness plays a more fundamental role than physicalism allows.

Science as a label is sometimes used to undermine alternative metaphysical positions in a way closer to scientism than to scientific observation. Philosophy has not resolved the metaphysical debate, but neither has science. The appearance of quantum support for physicalism is interpretation and supposition, not empirical verdict.

If consciousness has a non-physical component, then quantum theory is complete — the measurement problem has a solution. If consciousness is entirely physical, then quantum mechanics remains incomplete, the measurement problem unresolved. This is not a question science can answer through measurement alone. The phenomena themselves — entanglement, superposition, the observer effect — sit more naturally within frameworks that treat consciousness as fundamental rather than emergent. Physicalism and physics are not interchangeable, and the evidence, honestly assessed, appears to point beyond the purely physical.

About the Author

Sopa Thaye is a teacher of philosophy and religion in the UK and author of the philosophy Substack *The Primacy of Perception*. He is currently completing a book on epistemology, consciousness, and metaphysics.

Publication Information

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.